Reviewer 2

General comments:

This is an interesting paper comparing global leaf area index (LAI) simulations with satellite-derived LAI observations. Seven land surface models are considered along with 3 satellite LAI data sets. All products and observations are first projected onto the same 0.5 degree x 0.5 degree grid and monthly averages are considered. Maps of simulated and observed LAI peak, trough, and annual amplitude are presented. Mean monthly LAI time series show that the models tend to simulate LAI peaks later than the observations. The paper is reasonably well written, but the discussion section could be improved.

We thank the reviewer for her/his evaluation of the manuscript. We will address the discussion section in the revised version of the manuscript.

Recommendation: major revisions.

Particular comment:

- Phenology simulations may be affected by errors in the atmospheric forcing. What is the quality of the CRUNCEP atmospheric forcing database? Has it been evaluated in previous studies?

We agree with the reviewer that the atmospheric dataset used to force the land surface models has a relevant impact on the final simulated phenology. The selection of CRUNCEP derives from the TRENDY protocol used as a reference for the CRESCENDO experiments. More recent efforts apply other atmospheric forcing, such as GSWP3 or CRUJRA. We will add a short discussion on this point in the discussion section of the manuscript, even if the evaluation of atmospheric forcing is out of the scope of the present study since all LSMs are forced by the same atmospheric forcing here.

- I am concerned with the definition of "LAI trough". Unlike the LAI peak, low LAI values (e.g. in winter or during a drought) can persist for several months. Commonly used phenology indicators, in addition to peak time, are leaf onset and leaf offset. Why not use these more common indicators? Is LAI trough equivalent to leaf offset?

The present study expands the results presented in its companion paper (Peano et al., 2021), in which the leaf onset and offset timings simulated by the same seven LSMs are compared against the same set of satellite observations. The LAI trough evaluated in the present study represents the peak of dormancy. However, as pointed out by the reviewer, the trough may persist for more than one month. For this reason, we assume the LAI trough/dormancy month is the first month reaching the lowest value. We will ensure this point is clear and further discuss the implications of this assumption in the discussion section of the revised manuscript.

- The observed model LAI peak lags at mid and high latitudes may indicate a problem in the representation of temperature. Temperature is a key driver of leaf emergence in these regions. The temperature relevant to phenology is likely to be close to the land surface temperature resulting from the energy budget calculations. This is particularly true for the ISBA model, where phenology is driven by photosynthesis and leaf temperature. Is daytime leaf temperature underestimated in this model? In a number of ISBA papers (e.g. https://doi.org/10.5194/hess-21-4861-2017) the LAI peak time is generally consistent with observations. What has changed in the ISBA settings? Is the surface temperature calculated in the same way as before?

We agree with the reviewer on the high relevance of temperature as a key driver of leaf emergence, growth and decay. For this reason, as noted by the reviewer, the selected atmospheric forcing should be clearly accounted for as a possible source of discrepancy between modelled and observed timings. This point will be mentioned and addressed in the revised version of the discussion section.

Regarding ISBA results, the version used in the present study is based on the CMIP6 configuration described in Decharme et al. (2019) and Delire et al. (2020). As mentioned before, the difference in atmospheric forcing (i.e. CRUNCEP here and WFDEI in Dewaele et al., 2017) could partially explain the discrepancies between the two studies.

- For all models, it should be stated how the temperature used in the phenology model is calculated and how reliable it is.

Following the comments from the first reviewer, we will introduce a table in which the main drivers and approaches used in representing LAI and phenology in each LSM will be referenced as done in Table 1 in the companion paper (Peano et al., 2021).

References

Decharme, B., Delire, C., Minvielle, M., Colin, J., Vergnes, J.-P., Alias, A., Saint-Martin, D., Séférian, R., Sénési, S., and Voldoire, A.:Recent changes in the ISBA-CTRIP land surface system for use in the CNRM-CM6 climate model and in global off-line hydrological applications, Journal of Advances in Modeling Earth Systems, 11, 1207–1252, https://doi.org/10.1029/2018MS001545, 2019.

Delire, C., Séférian, R., Decharme, B., Alkama, R., Calvet, J.-C., Carrer, D., Gibelin, A.-L., Joetzjer, E., Morel, X., Rocher, M., and Tzanos, D.: The global land carbon cycle simulated with ISBA: improvements over the last decade, Journal of Advances in Modeling Earth Systems, 12, e2019MS001 886, https://doi.org/10.1029/2019MS001886, 2020.

Dewaele, H., Munier, S., Albergel, C., Planque, C., Laanaia, N., Carrer, D., and Calvet, J.-C.: Parameter optimisation for a better representation of drought by LSMs: inverse modelling vs. sequential data assimilation, Hydrol. Earth Syst. Sci., 21, 4861–4878, https://doi.org/10.5194/hess-21-4861-2017, 2017.

Peano, D., Hemming, D., Materia, S., Delire, C., Fan, Y., Joetzjer, E., Lee, H., Nabel, J. E. M. S., Park, T., Peylin, P., Wårlind, D., Wiltshire, A., and Zaehle, S.: Plant phenology evaluation of CRESCENDO land surface models – Part 1: Start and end of the growing season, Biogeosciences, 18, 2405–2428, https://doi.org/10.5194/bg-18-2405-2021, 2021.